0550
1011

⅃ S OIPE

```
RAW SEQUENCE LISTING              DATE: 11/27/2001
PATENT APPLICATION:  US/09/837,138    TIME: 10:44:23

     Input Set : N:\Crf3\RULE60\09837138.raw
     Output Set: N:\CRF3\11272001\I837138.raw
```

```
 1 <110> APPLICANT: Petrini, John H.J.
 2        Morgan, William Franklin
 3        Maser, Richard Scott
 4        Carney, James Patrick
 5 <120> TITLE OF INVENTION: DNA Encoding A DNA Repair Protein
 6 <130> FILE REFERENCE: 800.019US1
 7 <140> CURRENT APPLICATION NUMBER: 09/837,138
 8 <141> CURRENT FILING DATE: 2001-04-18
 9 <150> PRIOR APPLICATION NUMBER: 09/067,641
10 <151> PRIOR FILING DATE: 1999-12-03
11 <160> NUMBER OF SEQ ID NOS: 24
12 <170> SOFTWARE: FastSEQ for Windows Version 3.0
14 <210> SEQ_ID NO: 1
15 <211> LENGTH: 4403
16 <212> TYPE: DNA
17 <213> ORGANISM: Homo sapiens
18 <400> SEQUENCE: 1
19     ttcggcacga ggcgcggttg cacgtcggcc ccagccctga ggagccggac cgatgtggaa      60
20     actgctgccc gccgcgggcc cggcaggagg agaaccatac agacttttga ctggcgttga     120
21     gtacgttgtt ggaaggaaaa actgtgccat tctaattgaa aatgatcagt cgatcagccg     180
22     aaatcatgct gtgttaactg ctaacttttc tgtaaccaac ctgagtcaaa cagatgaaat     240
23     ccctgtattg acattaaaag ataattctaa gtatggtacc tttgttaatg aggaaaaaat     300
24     gcagaatggc ttttcccgaa cttttgaagtc gggggatggt attacttttg gagtgtttgg     360
25     aagtaaattc agaatagagt atgagccttt ggttgcatgc tcttcttgtt tagatgtctc     420
26     tgggaaaact gctttaaatc aagctatatt gcaacttgga ggatttactg taaacaattg     480
27     gacagaagaa tgcactcacc ttgtcatggt atcagtgaaa gttaccatta aaacaatatg     540
28     tgcactcatt tgtggacgtc caattgtaaa gccagaatat tttactgaat tcctgaaagc     600
29     agttcagtcc aagaagcagc ctccacaaat tgaaagtttt tacccacctc ttgatgaacc     660
30     atctattgga agtaaaaatg ttgatctgtc aggacggcag gaaagaaaac aaatcttcaa     720
31     agggaaaaca tttatatttt tgaatgccaa acagcataag aaattgagtt ccgcagttgt     780
32     ctttggaggt ggggaagcta ggttgataac agaagagaat gaagaagaac ataatttctt     840
33     tttggctccg ggaacgtgtg ttgttgatac aggaataaca aactcacaga ccttaattcc     900
34     tgactgtcag aagaaatgga ttcagtcaat aatggatatg ctccaaaggc aaggtcttag     960
35     acctattcct gaagcagaaa ttggattggc ggtgattttc atgactacaa agaattactg    1020
36     tgatcctcag ggccatccca gtacaggatt aaagacaaca actccaggac caagcctttc    1080
37     acaaggcgtg tcagttgatg aaaaactaat gccaagcgcc ccagtgaaca ctacaacata    1140
38     cgtagctgac acagaatcag agcaagcaga tacatgggat ttgagtgaaa ggccaaaaga    1200
39     aatcaaagtc tccaaaatgg aacaaaaatt cagaatgctt tcacaagacg cacccactgt    1260
40     aaaggagtcc tgcaaaacaa gctctaataa taatagtatg gtatcaaata ctttggctaa    1320
41     gatgagaatc ccaaactatc agctttcacc aactaaattg ccaagtataa ataaaagtaa    1380
42     agatagggct tctcagcagc agcagaccaa ctccatcaga aactactttc agccgtctac    1440
43     caaaaaaagg gaaagggatg aagaaaatca agaaatgtct tcatgcaaat cagcaagaat    1500
44     agaaacgtct tgttctcttt tagaacaaac acaacctgct acaccctcat gtgtggaaaaa    1560
45     taaggagcag catctatctg agaatgagcc tgtggacaca aactcagaca ataacttatt    1620
46     tacagataca gatttaaaat ctattgtgaa aaattctgcc agtaaatctc atgctgcaga    1680
47     aaagctaaga tcaaataaaa aaagggaaat ggatgatgtg gccatagaag atgaagtatt    1740
48     ggaacagtta ttcaaggaca caaaaccaga gttagaaatt gatgtgaaag ttcaaaaaca    1800
```

ENTERED

PATENT APPLICATION: US/09/837,138   TIME: 10:44:23

       Input Set : N:\Crf3\RULE60\09837138.raw
       Output Set: N:\CRF3\11272001\I837138.raw

```
49    ggaggaagat gtcaatgtta gaaaaaggcc aaggatggat atagaaacaa atgacacttt    1860
50    cagtgatgaa gcagtaccag aaagtagcaa aatatctcaa gaaaatgaaa ttgggaagaa    1920
51    acgtgaactc aaggaagact cactatggtc agctaaagaa atatctaaca atgacaaact    1980
52    tcaggatgat agtgagatgc ttccaaaaaa gctgttattg actgaattta gatcactggt    2040
53    gattaaaaac tctacttcca gaaatccgtc tggcataaat gatgattatg gtcaactaaa    2100
54    aaatttcaag aaattcaaaa aggtcacata tcctggagca ggaaaacttc cacacatcat    2160
55    tggaggatca gatctaatag ctcatcatgc tcgaaagaat acagaactag aagagtggct    2220
56    aaggcaggaa atggaggtac aaaatcaaca tgcaaaagaa gagtctcttg ctgatgatct    2280
57    ttttagatac aatccttatt taaaaaggag aagataactg aggattttaa aaagaagcca    2340
58    tggaaaaact tcctagtaag catctacttc aggccaacaa ggttatatga atatatagtg    2400
59    tatagaagcg atttaagtta caatgtttta tggcctaaat ttattaaata aaatgcacaa    2460
60    aactttgatt cttttgtatg taacaattgt ttgtyctgtt ttcaggcttt gtcattgcat    2520
61    ctttttttca tttttaaatg tgttttgttt attaaatagt taatatagtc acagttcaaa    2580
62    attctaaatr tacgtaaggt aaaggactaa agtcaccctt ccaccattgt cctagctact    2640
63    tattttttaaa taatttccta cacaaatgat agcataacat atgcagtgtt ctacaccttg    2700
64    ctttttttact tagtaagatt aaaaattata ggaatatcaa tataatgttt ttaatatttt    2760
65    ttcttttcca ttatgctgta gtcttaccta aactctggtg atccaaacaa aatggcttca    2820
66    gtggtgcaga tgtcacctac atgttattct agtactagaa actgaagacc atgtggagac    2880
67    ttcatcaaac atgggtttag ttttcaccag aatggaaaga cctgtacccc tttttggtgg    2940
68    tcttactgag ctgggtgggt gtctgttttg agcttattta gagtcctagt tttcctactt    3000
69    ataaagtaga aatggtgaga ttgttttctt tttctacckt aaaggggagat ggtaagaaac    3060
70    aatgaatgtc tttttttcaaa ctttattgac aagtgatttt caagtctgtg ttcaaaaata    3120
71    tattcatgta cctgtgatcc agcaagaagg gagttccagt caagagtcac tacaactgat    3180
72    tagttgttta gagaatgaga aatggaacag tgaggaatgg aggccatatt tccatgactt    3240
73    cccttgtaaa cagaagcaac agaagggaca agaggctggc ctctacatca ctctcacctt    3300
74    ccaaatcttg tggaagtgca tctacttgcc agaaccaaat taacttactt ccaagttctg    3360
75    gctgcttgca ggtggaactc cagctgcaag ggagttaggg aaatgaaggt ctttttttaa    3420
76    aagcttctca gccttcctag ggaacagaaa ttgggtgagc caatctgcaa tttctactac    3480
77    aggcattgag accagttaga ttattgaaat attatagaga gttatgaaca cttaaattat    3540
78    gatagtggta tgacattgga tagaacatgg gatactttag aagtagaatt gacaggcat    3600
79    attagttgat gaaatggagt catttgagtc tyttaatagc catgtatcat aattaccaag    3660
80    tgaagctggt ggaacatatg gtctccattt tacagttaag gaatataatg gacagattaa    3720
81    tattgttytc tgtcatgccc acaatccctt tctaaggaag actgccctac tatagcagtt    3780
82    tttatatttg tcaatttatg aatataatga atgaggagtt ctggtacctc ctgtctttac    3840
83    aaatattggg tgttgtccag tattttttccc tttttaaccm ttcccaattc gggtgtgtag    3900
84    gtggatgttt ccatttgggt tttaatttgt atatccctga tagctataat tgggtcatag    3960
85    aaattcttta tacattctag atgcaagtct cttgycggat atacgtattg agatattaca    4020
86    cctagtctgt ggcttgactg ttttctttat gtcttttgat gaatagaagt tttaaatttt    4080
87    gacaaggtca aatttatttt tttcttttgt ttgatatttt ttctctccaa tttaaccccca    4140
88    agatttcaga tattctgctc tattatataa actttatatt tttatatttg tgatctacct    4200
89    tgaattgata tgtatgttgt gaattatgga tcagggttct tttttttcccc catacaagta    4260
90    tccagtcatt gtaacactgt ttattgaaag aattatcctt tcctcattaa attaccttgc    4320
91    caattagtaa aaaatcaatt aaccatrmar mmmrrrggat ccactagttc tagagcggcc    4380
92    gccaccgcgg tggagctcca gct                                             4403
94  <210> SEQ ID NO: 2
95  <211> LENGTH: 754
96  <212> TYPE: PRT
97  <213> ORGANISM: Homo sapiens
98  <400> SEQUENCE: 2
```

RAW SEQUENCE LISTING   DATE: 11/27/2001
PATENT APPLICATION: US/09/837,138   TIME: 10:44:23

Input Set : N:\Crf3\RULE60\09837138.raw
Output Set: N:\CRF3\11272001\I837138.raw

99	Met	Trp	Lys	Leu	Leu	Pro	Ala	Ala	Gly	Pro	Ala	Gly	Gly	Glu	Pro	Tyr
100	1				5					10					15
101	Arg	Leu	Leu	Thr	Gly	Val	Glu	Tyr	Val	Val	Gly	Arg	Lys	Asn	Cys	Ala
102				20				25					30
103	Ile	Leu	Ile	Glu	Asn	Asp	Gln	Ser	Ile	Ser	Arg	Asn	His	Ala	Val	Leu
104			35				40					45
105	Thr	Ala	Asn	Phe	Ser	Val	Thr	Asn	Leu	Ser	Gln	Thr	Asp	Glu	Ile	Pro
106		50				55					60
107	Val	Leu	Thr	Leu	Lys	Asp	Asn	Ser	Lys	Tyr	Gly	Thr	Phe	Val	Asn	Glu
108	65				70				75					80
109	Glu	Lys	Met	Gln	Asn	Gly	Phe	Ser	Arg	Thr	Leu	Lys	Ser	Gly	Asp	Gly
110				85				90				95
111	Ile	Thr	Phe	Gly	Val	Phe	Gly	Ser	Lys	Phe	Arg	Ile·Glu	Tyr	Glu	Pro
112			100				105					110
113	Leu	Val	Ala	Cys	Ser	Ser	Cys	Leu	Asp	Val	Ser	Gly	Lys	Thr	Ala	Leu
114			115				120				125
115	Asn	Gln	Ala	Ile	Leu	Gln	Leu	Gly	Gly	Phe	Thr	Val	Asn	Asn	Trp	Thr
116		130				135				140
117	Glu	Glu	Cys	Thr	His	Leu	Val	Met	Val	Ser	Val	Lys	Val	Thr	Ile	Lys
118	145				150				155					160
119	Thr	Ile	Cys	Ala	Leu	Ile	Cys	Gly	Arg	Pro	Ile	Val	Lys	Pro	Glu	Tyr
120				165				170					175
121	Phe	Thr	Glu	Phe	Leu	Lys	Ala	Val	Gln	Ser	Lys	Lys	Gln	Pro	Pro	Gln
122			180				185					190
123	Ile	Glu	Ser	Phe	Tyr	Pro	Pro	Leu	Asp	Glu	Pro	Ser	Ile	Gly	Ser	Lys
124			195				200				205
125	Asn	Val	Asp	Leu	Ser	Gly	Arg	Gln	Glu	Arg	Lys	Gln	Ile	Phe	Lys	Gly
126		210				215				220
127	Lys	Thr	Phe	Ile	Phe	Leu	Asn	Ala	Lys	Gln	His	Lys	Lys	Leu	Ser	Ser
128	225				230				235					240
129	Ala	Val	Val	Phe	Gly	Gly	Gly	Glu	Ala	Arg	Leu	Ile	Thr	Glu	Glu	Asn
130				245				250				255
131	Glu	Glu	Glu	His	Asn	Phe	Phe	Leu	Ala	Pro	Gly	Thr	Cys	Val	Val	Asp
132			260				265				270
133	Thr	Gly	Ile	Thr	Asn	Ser	Gln	Thr	Leu	Ile	Pro	Asp	Cys	Gln	Lys	Lys
134			275				280				285
135	Trp	Ile	Gln	Ser	Ile	Met	Asp	Met	Leu	Gln	Arg	Gln	Gly	Leu	Arg	Pro
136		290				295				300
137	Ile	Pro	Glu	Ala	Glu	Ile	Gly	Leu	Ala	Val	Ile	Phe	Met	Thr	Thr	Lys
138	305				310				315					320
139	Asn	Tyr	Cys	Asp	Pro	Gln	Gly	His	Pro	Ser	Thr	Gly	Leu	Lys	Thr	Thr
140				325				330				335
141	Thr	Pro	Gly	Pro	Ser	Leu	Ser	Gln	Gly	Val	Ser	Val	Asp	Glu	Lys	Leu
142			340				345				350
143	Met	Pro	Ser	Ala	Pro	Val	Asn	Thr	Thr	Thr	Tyr	Val	Ala	Asp	Thr	Glu
144			355				360				365
145	Ser	Glu	Gln	Ala	Asp	Thr	Trp	Asp	Leu	Ser	Glu	Arg	Pro	Lys	Glu	Ile
146			370				375				380
147	Lys	Val	Ser	Lys	Met	Glu	Gln	Lys	Phe	Arg	Met	Leu	Ser	Gln	Asp	Ala

RAW SEQUENCE LISTING   DATE: 11/27/2001
PATENT APPLICATION: US/09/837,138   TIME: 10:44:23

Input Set : N:\Crf3\RULE60\09837138.raw
Output Set: N:\CRF3\11272001\I837138.raw

```
148        385                   390                   395                   400
149        Pro Thr Val Lys Glu Ser Cys Lys Thr Ser Ser Asn Asn Asn Ser Met
150                        405                   410                   415
151        Val Ser Asn Thr Leu Ala Lys Met Arg Ile Pro Asn Tyr Gln Leu Ser
152                    .   420                   425                   430
153        Pro Thr Lys Leu Pro Ser Ile Asn Lys Ser Lys Asp Arg Ala Ser Gln
154                        435                   440                   445
155        Gln Gln Gln Thr Asn Ser Ile Arg Asn Tyr Phe Gln Pro Ser Thr Lys
156             450                   455                   460
157        Lys Arg Glu Arg Asp Glu Glu Asn Gln Glu Met Ser Ser Cys Lys Ser
158        465                   470                   475                   480
159    .   Ala Arg Ile Glu Thr Ser Cys Ser Leu Leu Glu Gln Thr Gln Pro Ala
160                        485                   490                   495
161        Thr Pro Ser Leu Trp Lys Asn Lys Glu Gln His Leu Ser Glu Asn Glu
162                        500                   505                   510
163        Pro Val Asp Thr Asn Ser Asp Asn Asn Leu Phe Thr Asp Thr Asp Leu
164             515                   520                   525
165        Lys Ser Ile Val Lys Asn Ser Ala Ser Lys Ser His Ala Ala Glu Lys
166             530                   535         ,         540
167        Leu Arg Ser Asn Lys Lys Arg Glu Met Asp Asp Val Ala Ile Glu Asp
168        545                   550                   555                   560
169        Glu Val Leu Glu Gln Leu Phe Lys Asp Thr Lys Pro Glu Leu Glu Ile
170                        565                   570                   575
171        Asp Val Lys Val Gln Lys Gln Glu Glu Asp Val Asn Val Arg Lys Arg
172                        580                   585                   590
173        Pro Arg Met Asp Ile Glu Thr Asn Asp Thr Phe Ser Asp Glu Ala Val
174                        595                   600                   605
175        Pro Glu Ser Ser Lys Ile Ser Gln Glu Asn Glu Ile Gly Lys Lys Arg
176             610                   615                   620
177        Glu Leu Lys Glu Asp Ser Leu Trp Ser Ala Lys Glu Ile Ser Asn Asn
178        625                   630                   635                   640
179        Asp Lys Leu Gln Asp Asp Ser Glu Met Leu Pro Lys Lys Leu Leu Leu
180                        645                   650                   655
181        Thr Glu Phe Arg Ser Leu Val Ile Lys Asn Ser Thr Ser Arg Asn Pro
182                        660                   665                   670
183        Ser Gly Ile Asn Asp Asp Tyr Gly Gln Leu Lys Asn Phe Lys Lys Phe
184                        675                   680                   685
185        Lys Lys Val Thr Tyr Pro Gly Ala Gly Lys Leu Pro His Ile Ile Gly
186             690                   695                   700
187        Gly Ser Asp Leu Ile Ala His His Ala Arg Lys Asn Thr Glu Leu Glu
188        705                   710                   715                   720
189        Glu Trp Leu Arg Gln Glu Met Glu Val Gln Asn Gln His Ala Lys Glu
190                        725                   730                   735
191        Glu Ser Leu Ala Asp Asp Leu Phe Arg Tyr Asn Pro Tyr Leu Lys Arg
192                        740                   745                   750
193        Arg Arg
195 <210> SEQ ID NO: 3
196 <211> LENGTH: 87
197 <212> TYPE: PRT
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/837,138

DATE: 11/27/2001
TIME: 10:44:23

Input Set : N:\Crf3\RULE60\09837138.raw
Output Set: N:\CRF3\11272001\I837138.raw

```
198 <213> ORGANISM: Homo sapiens
199 <220> FEATURE:
200 <221> NAME/KEY: UNSURE
201 <222> LOCATION: (48)...(48)
202 <223> OTHER INFORMATION: Unsure
203 <400> SEQUENCE: 3
204        Tyr Val Val Gly Arg Lys Asn Cys Ala Ile Leu Ile Glu Asn Asp Gln
205        1               5                   10                  15
206        Ser Ile Ser Arg Asn His Ala Val Leu Thr Ala Asn Phe Ser Val Thr
207                    20                  25                  30
W--> 208   Asn Leu Ser Gln Thr Asp Glu Ile Pro Val Leu Thr Leu Lys Asn Xaa
209                35                  40                  45
210        Lys Tyr Gly Thr Phe Val Asn Glu Glu Lys Met Gln Asn Gly Phe Ser
211            50                  55                  60
212        Arg Thr Leu Lys Ser Val Asp Gly Ile Thr Phe Gly Val Phe Gly Ser
213        65                  70                  75                  80
214        Lys Phe Arg Ile Glu Tyr Glu
215                        85
217 <210> SEQ ID NO: 4
218 <211> LENGTH: 87
219 <212> TYPE: PRT
220 <213> ORGANISM: Homo sapiens
221 <400> SEQUENCE: 4
222        Tyr Ser Ile Gly Arg Ser Ser Lys Asn Pro Leu Ile Ile Lys Asn Asp
223        1               5                   10                  15
224        Lys Ser Ile Ser Arg Gln His Ile Thr Phe Lys Trp Glu Ile Asn Asn
225                    20                  25                  30
226        Ser Ser Asp Leu Lys His Ser Ser Leu Cys Leu Val Asn Lys Gly Lys
227                    35                  40                  45
228        Leu Thr Ser Leu Asn Lys Lys Phe Met Lys Val Gly Glu Thr Phe Thr
229                50                  55                  60
230        Ile Asn Ala Ser Cys Val Leu Lys Ser Thr Ile Glu Leu Gly Thr Thr
231        65                  70                  75                  80
232        Pro Ile Arg Ile Glu Phe Glu
233                        85
235 <210> SEQ ID NO: 5
236 <211> LENGTH: 13
237 <212> TYPE: PRT
238 <213> ORGANISM: Homo sapiens
239 <400> SEQUENCE: 5
240        Asn Pro Ser Gly Leu Asn Asp Asp Tyr Gly Gln Leu Lys
241        1               5                   10
243 <210> SEQ ID NO: 6
244 <211> LENGTH: 680
245 <212> TYPE: PRT
246 <213> ORGANISM: Homo sapiens
247 <400> SEQUENCE: 6
248        Met Ser Thr Ala Asp Ala Leu Asp Asp Glu Asn Thr Phe Lys Ile Leu
249        1               5                   10                  15
```

VERIFICATION SUMMARY   DATE: 11/27/2001
PATENT APPLICATION: US/09/837,138   TIME: 10:44:24

Input Set : N:\Crf3\RULE60\09837138.raw
Output Set: N:\CRF3\11272001\I837138.raw

L:208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3